

Analisis Data Rawak Terikan-Lesu Menggunakan Kaedah Statistik Purata-Tergerak

Shahrum Abdullah, Mohammad Darahim Ibrahim, Azami Zaharim dan Zulkifli Mohd Nopiah

Fakulti Kejuruteraan,
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor,
Malaysia
E-mail: Shahrum@vlsi.eng.ukm.my

Received Date: 15th May 2007 Accepted Date: 17th February 2009

ABSTRAK

Kajian ini dijalankan untuk menganalisis data terikan-lesu yang diukur pada sistem ampaian sebuah kereta. Oleh itu, teknik ini diperkenalkan bagi menganalisis kebarangkalian serta mendasari data yang berkaitan dengan beban terikan-lesu. Dalam analisis lesu bahan dan komponen, kerosakan lesu biasanya disebabkan oleh kesan daripada satu pembolehubah rawak, yang biasanya dibentuk daripada satu set data siri-masa. Menerusi data siri-masa ini, kajian data lesu dapat diperkembangkan menerusi kaedah statistik dan model analisis yang dicadangkan ialah model purata-tergerak. Model ini dapat menganggarkan trend purata-tergerak bagi mengurangkan variasi rawak dalam suatu data. Dalam kajian ini, data terikan-lesu telah diperolehi daripada pengukuran ke atas komponen sistem ampaian kereta yang dipandu di atas permukaan jalanraya luar kawasan perbandaran. Data ini diukur pada frekuensi persampelan 200 Hz dalam tempoh 300 saat. Kemudian, data ini telah dilakukan analisis purata-tergerak menerusi kaedah Peratusan Ralat Min Mutlak (MAPE), Sisihan Min Mutlak (MAD) dan Sisihan Min Kuasa Dua (MSD). Analisis korelasi data berdasarkan kepada Fungsi Autokorelasi (ACF) dan Fungsi Separa Autokorelasi (PACF) juga dibincangkan. Secara kesimpulannya, dapat dirumuskan bahawa kaedah-kaedah ini dapat digunakan untuk memproses data rawak yang berubah-ubah bagi penghasilan satu data yang pegun, dengan menyingkirkan bahagian-bahagian data yang tidak pegun.

Kata kunci: Data rawak, lesu, purata-tergerak, siri-masa, statistik.

ABSTRACT

This paper presents the analysis of fatigue road loading which was measured on a component of an automobile suspension system. Thus, this significant technique was introduced for preserving data associated to the underlying probabilistic properties that are related to the fatigue damage. The analysis was based on the fatigue damage potential, which was related to the random variables in a time series data. Using this data type, the analysis was performed by means of the statistical method, such as the Moving Average technique. This model can then be applied to estimate the data trend by using the related Moving Average models in order

to reduce a random variation of fatigue data. In this study, the data were experimentally measured on an automobile suspension system which was travelling over a public road surface. During the measurement, the data collection was performed at the sampling rate of 200Hz in the 300 second of the record length. For the analysis, the Moving Average method was applied to the fatigue data in order to determine the parameters related to Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD) and Mean Squared Deviation (MSD). In addition, the significant analysis of data correlation based on Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) was also discussed. Finally, it is suggested that this method provided a good platform to process the changeable random fatigue strain loading in order to produce a stationary data, with the removal of the nonstationary parts.

Keywords: Random data, fatigue, moving-average, time series, statistic.

PENDAHULUAN

Isyarat adalah satu siri nombor yang diperolehi daripada pengukuran dan kebiasaannya ia diperolehi dengan menggunakan kaedah pengukuran yang ditentukan sebagai satu fungsi (Meyer 1993).

Salah satu data siri masa ialah data lesu yang diukur ke atas suatu komponen yang dikenakan daya berulang, dan parameter yang biasa digunakan ialah dalam bentuk terikan. Dalam kes siri masa, ia mengandungi satu set pembolehubah pemerhatian di mana ianya telah diukur dalam satu jangka masa tertentu (Harvey 1981). Ketika ini, kebanyakan ukuran yang diperolehi adalah daripada ujikaji, dan data berkenaan direkodkan secara digital. Rekod data ini dikenali sebagai satu siri masa diskret. Hasil daripada analisis ini digunakan untuk menentukan ciri-ciri fungsi asal secara statistik dengan memanipulasi siri data diskret (Newland 1993).

Kebanyakan aplikasi analisis data yang biasa digunakan ialah berkaitan dengan data yang privasi-sensitif. Antara cara terbaik untuk mendapatkan angka-angka rawak yang tidak dijangka ialah dengan mengukur fenomena fizik yang berlaku seperti lesu, pereputan radioaktif, hingga terma dalam semikonduktor dan malah imej digital bagi sebuah lampu lava. Walau bagaimanapun terdapat beberapa komputer (atau pengguna) perlu mengakses beberapa jenis perisian yang khusus bagi memproses data rawak ini (Yan et al. 2001).

Berasaskan sifat-sifat data siri masa, justeru, objektif kajian ini adalah untuk mengkaji satu teknik yang sesuai dalam menganalisis data rawak lesu (dalam bentuk terikan satu-paksi) dengan menggunakan kaedah purata-tergerak. Secara tidak langsung, ia menjurus kepada pembangunan satu kelas algoritma data (Agrawal and Srikant 2000; Agrawal and

Aggawal 2001) dan algoritma pengkelasan data ini juga dapat digunakan dalam pendekatan mengekstrak bahagian data yang memberi kesan kepada kerosakan lesu. Pendekatan ini juga boleh digunakan untuk mengekalkan perlakuan asal data berkenaan daripada gangguan persekitaran yang mungkin berlaku (Estivill-Castro and Brankovik 1999). Lazimnya, data ini telah digunakan dengan teknik-teknik penyuaian lengkung untuk membangunkan tingkah laku lesu bahan di atas satu julat yang sesuai pada tahap-tahap tekanan yang berbeza.

LATAR BELAKANG KAJIAN

Secara amnya, kebanyakan isyarat mempunyai sifat rawak atau ciri-ciri yang kurang jelas, justeru penakrifan lengkap bagi memahami maklumat siri data ini adalah diperlukan. Ini merupakan satu cabaran dalam menganalisis data yang telah diperolehi daripada ujikaji. Salah satu jalan penyelesaian analisis data ini ialah menggunakan kaedah analisis statistik dan pemprosesan isyarat (Natrella 1966). Satu isyarat yang mewakili keadaan fizikal secara rawak tidak boleh digambarkan dengan satu kaedah di mana ia merupakan satu persamaan matematik yang berketentuan. Satu isyarat mewakili satu fenomena secara rawak di mana ia disifatkan sebagai pegun atau tidak bergerak.

Isyarat yang pegun disifatkan oleh nilai-nilai parameter statistik isyarat, seperti nilai min, nilai sisihan piawai dan nilai punca min kuasa dua, tidak banyak berubah apabila ia melintasi keseluruhan masa sesuatu isyarat. Proses-proses rawak secara pegun dapat dikategorikan sebagai ergodik atau bukan ergodik. Jika proses rawak adalah pegun, dan nilai min serta fungsi autokorelasi tidak berbeza apabila dianalisis pada segmen sampel yang berbeza, maka proses

rawak tersebut ditakrifkan sebagai ergodik. Dalam keadaan tidak pegun, nilai-nilai statistik isyaratnya adalah bergantung pada masa yang diambil (Tacer and Laughlin 1998). Isyarat tidak pegun dapat dibahagikan kepada dua kategori iaitu tidak pegun secara lemah dan tidak pegun secara kuat.

Umumnya, isyarat statistik digunakan untuk mengelaskan isyarat secara rawak. Parameter statistik yang biasa digunakan ialah nilai min, nilai sisihan piawai, nilai punca min kuasa dua (r.m.s), kepencongan kurtosis dan faktor puncak (Bendat and Piersol 1986). Untuk satu bilangan isyarat yang diwakili sebanyak n bilangan data diskret, nilai min \bar{x} ditakrifkan secara matematik menerusi persamaan berikut:

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j \quad (1)$$

Selain daripada itu, nilai sisihan piawai (SD) ditakrifkan menerusi persamaan matematik berikut:

$$SD = \left\{ \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2 \right\}^{\frac{1}{2}} \quad (2)$$

bagi sampel yang lebih daripada 30 (Bendat and Piersol 1986). Nilai sisihan piawai diperolehi daripada perbezaan antara data asal dengan nilai min.

Dalam analisis statistik kejuruteraan, nilai punca-min-kuasa dua (*root-mean-square - r.m.s*) adalah isyarat momen statistik yang ke-2 dan ia mewakili satu penjumlahan kandungan tenaga getaran isyarat dalam domain masa. Untuk isyarat yang nilai puratanya adalah sifar, nilai r.m.s. untuk isyarat ini adalah bersamaan dengan nilai sisihan piawai. Untuk data set diskret, nilai r.m.s diberikan menerusi persamaan matematik berikut:

$$r.m.s. = \left\{ \frac{1}{n} \sum_{j=1}^n X_j^2 \right\}^{\frac{1}{2}} \quad (3)$$

Kepencongan, ialah satu parameter momen statistik isyarat yang ke 3, di mana ia merupakan ukuran simetri bagi pengedaran titik-titik data secara simetri dengan nilai min. Kepencongan merupakan satu taburan simetri bagi satu isyarat sinusoid. Untuk satu isyarat rawak Gaussian, nilai kepencongannya adalah sifar. Nilai kepencongan negatif menunjukkan bahawa kebarangkalian taburan data untuk bergerak ke kiri manakala nilai

yang positif memberikan nilai kebarangkaliannya bergerak ke kanan. Persamaan matematik bagi nilai kepencongan adalah seperti berikut:

$$S = \frac{1}{n(SD)^3} \sum_{j=1}^n (X_j - \bar{X})^3 \quad (4)$$

Kurtosis ialah parameter momen statistik yang ke-4 dan ia merupakan satu parameter statistik yang peka terhadap data. Bagi data set yang diskret, nilai kurtosis dapat diperolehi menerusi persamaan berikut:

$$K = \frac{1}{n(r.m.s.)^4} \sum_{j=1}^n (X_j - \bar{X})^4 \quad (5)$$

Bagi satu taburan Gaussian, nilai kurtosis adalah lebih kurang 3.0. Nilai kurtosis yang lebih tinggi menunjukkan bahawa terdapat nilai amplitud yang tinggi dalam taburan Gaussian. Dalam bidang kejuruteraan, nilai kurtosis digunakan untuk mengesan sebarang kerosakan yang disebabkan oleh nilai-nilai beramplitud tinggi dalam bahagian isyarat (Hinton 1995).

Oleh kerana data yang tidak pegun dapat digunakan untuk menggambarkan corak secara rawak, purata-tergerak diperkenalkan sebagai satu pendekatan bagi mengurangkan variasi terhadap data siri masa (Qu and He 1986). Pelbagai kaedah digunakan untuk mengurangkan variasi dalam isyarat, tetapi ia perlu dimulakan dengan kaedah yang paling mudah kemudian diikuti dengan kaedah yang paling susah. Konsep ini dikenali sebagai kaedah purata-tergerak (Moving Average - MA). Terdapat beberapa kaedah purata-tergerak, antaranya ialah purata-tergerak mudah, purata-tergerak berganda dan purata-tergerak berpemberat (Makridakis et al. 1998). Model purata-tergerak ialah perhubungan yang saling bergantung antara terma-terma kesilapan secara berturut-turut. Model tersebut digunakan dalam mengira variasi yang terdahulu dan ia berasaskan kepada hubungan berikut:

$$Y_t = \theta_0 + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} + e_t \quad (6)$$

yang mana Y_t ialah nilai data asal, θ_0 ialah nilai yang berterusan, $\theta_1, \theta_2, \dots, \theta_{t-q}$ ialah pekali bagi linear regresi dan e_t ialah nilai pekali ralat.

Idea konsep purata-tergerak dapat digunakan kerana ia berkaitan dengan data yang mungkin wujud di mana nilainya menghampiri pada nilai data asal. Dengan mengambil nilai purata, ia

dapat menyediakan satu anggaran terhadap semua data dengan munasabah. Oleh itu, ia dapat menyingkirkan data yang tidak berkaitan dan juga dapat menghasilkan arah pergerakan yang baik bagi data asal yang tidak pegun (Holt et al. 1960). Kaedah purata-tergerak memerlukan bilangan pemerhatian yang terdiri dari bilangan sampel yang ganjil untuk menjadi pemberat bagi memperoleh nilai setiap purata (Kendall et al. 1983). Tujuan untuk keperluan ini bagi memastikan supaya purata berada di bahagian tengah (T_t) antara data asal diperolehi, keadaan tersebut diterjemahkan melalui persamaan (7):

$$T_t = \frac{1}{k} \sum_{j=-m}^m Y_{t+j} \quad (7)$$

yang mana nilai k ialah integer ganjil dan $m=(k-1)/2$. Purata-tergerak berganda merupakan teknik yang kedua yang digunakan dalam mengurangkan variasi terhadap data. Sebenarnya, sebarang kombinasi purata-tergerak dapat digunakan serentak dalam membentuk purata-tergerak berganda.

Purata-tergerak berpemberat ialah kes istimewa yang menyatakan $1/k$ sebagai pemberat bagi purata-tergerak k dan nilai purata ini ditulis sebagai

$$T_t = \sum_{j=-m}^m a_j Y_{t+j} \quad (8)$$

yang mana a_j ialah nilai faktor pemberat bagi kaedah purata-tergerak.

KAEDAH ANALISIS PURATA-TERGERAK

Kaedah purata-tergerak pernah digunakan keatas data lesu bagi menentukan model yang terbaik berdasarkan tiga kriteria, seperti yang ditunjukkan dalam Rajah 1 (Harvey 1981).

Dalam rajah ini, nilai Peratusan Ralat Min Mutlak (MAPE) bermaksud jumlah peratusan kesilapan-kesilapan untuk satu set data tanpa mengambil kira nilai positif dan negatif. Ia merupakan langkah yang tepat bagi kaedah kuantitatif (Qu and He 1986). Dalam kaedah MAPE ini, parameter PE_t adalah nilai relatif atau peratusan kesilapan bagi perbezaan di antara nilai asal dan nilai yang dijangkakan dalam analisis. PE_t bagi MAPE secara matematiknya diperolehi daripada

$$PE_t = \left(\frac{Y_t - F_t}{Y_t} \right) \times 100 \quad (9)$$

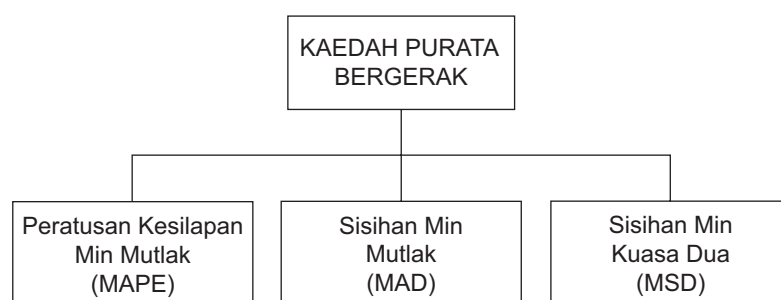
$$MAPE = \frac{1}{n} \sum_{t=1}^n |PE_t| \quad (10)$$

Sisihan Min Mutlak (MAD) ialah sisihan mutlak purata dari nilai min dan ianya biasa digunakan untuk mengunjurkan kesilapan yang berlaku dalam analisis. Perkataan sisihan min yang digunakan membawa maksud yang sama dengan Sisihan Min Mutlak, tetapi ianya tidak sama dari segi takrifan (dikenali sebagai operasi nilai mutlak), nilai sisihan min bagi semua data memberi nilai sifar (Holt et al. 1960; Kendall et al. 1983; Qu and He 1986; Makridakis et al. 1998). Secara hubungan matematik, parameter MAD ditakrifkan sebagai perkaitan berikut:

$$MAD = \frac{1}{n} \sum |Y_i - \bar{Y}| \quad (11)$$

Sisihan Min Kuasa Dua (MSD) ialah sisihan yang diperolehi dengan menjumlahkan sisihan kuasa dua dan dibahagikan dengan jumlah sampel data (Qu and He 1986). Persamaan matematiknya dinyatakan sebagai:

$$MSD = \frac{1}{n} \sum |Y_i - \bar{Y}|^2 \quad (12)$$



Rajah 1. Kriteria untuk menentukan mod purata-tergerak yang terbaik

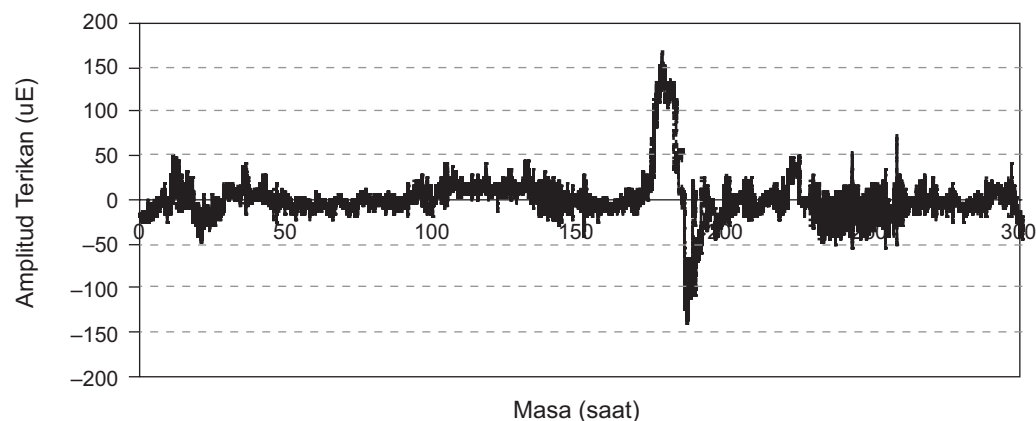
Dalam kajian ini, konsep terakhir yang digunakan ialah Fungsi Autokorelasi (ACF) dan Fungsi Separa Autokorelasi (PACF). Kaedah ACF biasanya digunakan untuk mengenalpasti data yang berturutan di mana ia dijelaskan pada situasi yang tertentu dalam menentukan data tersebut adalah tidak pegun. Tambahan lagi, ACF biasanya digunakan untuk mendapatkan model yang sesuai bagi data yang pegun. Sementara itu, pendekatan PACF digunakan untuk memperolehi perhubungan antara setiap sampel bagi data asal dan data semasa yang telah diproses. Oleh itu, fungsi autokorelasi adalah merujuk kepada kesesuaian pembolehubah yang digunakan dalam membuat analisis data siri masa.

KEPUTUSAN DAN PERBINCANGAN

Data yang telah diplot dalam Rajah 2 adalah satu data siri-masa yang diukur pada salah satu komponen sistem ampaian hadapan sebuah kereta. Semasa pengukuran data ini dijalankan, kereta berkenaan telah dipandu di atas satu permukaan jalanraya di luar kawasan

perbandaran. Amnya, jalanraya ini mempunyai permukaan berlubang, kasar serta biasa. Semasa pengukuran, data lesu (dalam unit mikroterikan) tersebut diperolehi pada kadar frekuensi persampelan 200 Hz dalam tempoh 300 saat. Justeru, bilangan data yang terhasil dalam tempoh proses pengukuran ini ialah sebanyak 60000 titik diskret.

Analisis pertama yang dilakukan ke atas data yang diplotkan dalam Rajah 2 ialah analisis statistik. Tujuan analisis ini dilakukan ialah untuk mentakrifkan kelakuan data siri-masa ini. Parameter statistik yang digunakan dalam analisis ini diterjemahkan pada Jadual 1. Dalam jadual ini, nilai punca-min-kuasa dua yang diperolehi (gambaran sebagai tenaga getaran isyarat dalam domain masa) ialah 0.104 mikroterikan. Nilai kurtosis data ini ialah 13.743 mikroterikan, dan nilai ini menunjukkan kewujudan perlakuan data yang tidak pegun. Tambahan lagi, data lesu yang diperbaiki ini memberikan nilai minimum ialah -138.72 mikroterikan dan nilai maksimum ialah 167.13 mikroterikan.



Rajah 2. Plot siri-masa untuk data terikan lesu pelbagai amplitud yang diukur pada komponen sistem ampaian kereta

Jadual 1. Nilai-nilai statistik data siri-masa terikan lesu yang diukur pada komponen sistem ampaian kereta

Keterangan		Nilai Statistik
Min [μE]		2.337
Sisihan Piawai [μE]		25.466
Minimum [μE]		-138.720
Maksimum [μE]		167.130
Kepencongan	Statistik	1.869
	Ralat Piawai	0.010
Kurtosis	Statistik	13.745
	Ralat Piawai	0.020

Kaedah analisis statistik yang lebih terperinci digunakan berasaskan kepada konsep purata-tergerak, Persamaan (13) hingga (15) digunakan untuk menghasilkan data yang pegun bagi pembolehubah amplitud data terikan lesu. Persamaan ini diperolehi dengan menggunakan satu perisian statistik iaitu SPSS. Persamaan dibawah dihasilkan dengan menganalisis data MA (1), MA (2) dan MA (3). Kemudian, kaedah purata-tergerak digunakan ke atas data asal ini bagi menghasilkan corak data yang pegun.

$$Y_1 = 2.336 + 0.935e_{t-1} + e_t \quad (13)$$

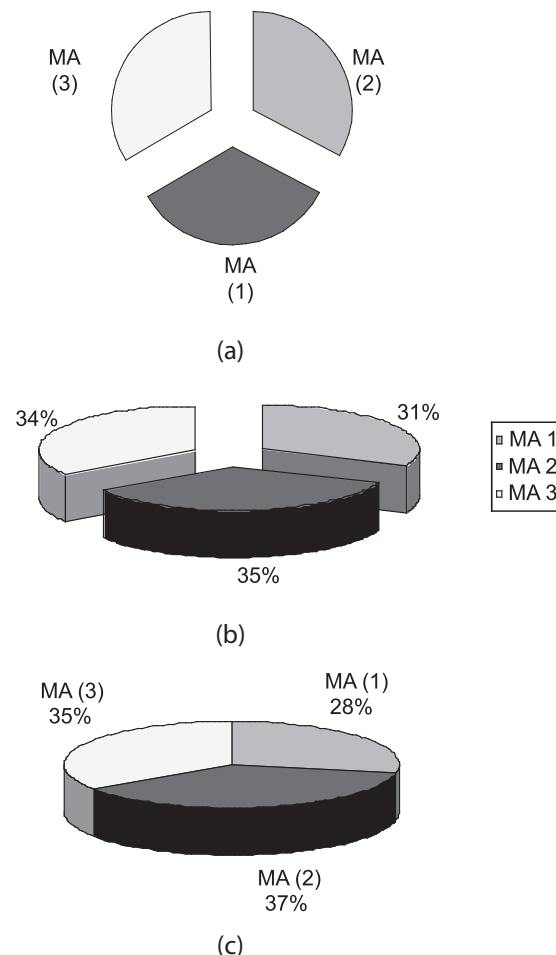
$$Y_2 = 2.335 + 0.498e_{t-1} + 0.920e_{t-2} + e_t \quad (14)$$

$$Y_3 = 2.335 + 0.325e_{t-1} + 0.564e_{t-2} + 0.801e_{t-3} + e_t \quad (15)$$

Kriteria nilai MAPE digunakan untuk menunjukkan bahawa MA(1) dan MA(2) mempunyai nilai ralat yang lebih kecil berbanding MA(3), iaitu sebanyak 7.36 mikroterikan (33%).

Nilai MA(1), MA(2) dan MA(3) masing-masing diperolehi berdasarkan kepada persamaan-persamaan (13), (14) dan (15). Keputusan yang diperolehi ini juga ditunjukkan menerusi carta pai dalam Rajah 3a. Dengan kaedah MAD yang ditunjukkan, MA(1) juga memberi nilai ralat yang lebih kecil berbanding MA(2) dan MA(3), iaitu 1.05 mikroterikan (31%).

Taburan nilai ralat ini ditunjukkan dalam Rajah 3b. Akhirnya, data siri-masa terikan lesu dianalisis dengan menggunakan kaedah MSD, dan ia menunjukkan bahawa MA(1) sekali lagi memberi nilai kesilapan yang terkecil berbanding dengan MA(2) dan MA(3). Berdasarkan analisis ini, ia menunjukkan bahawa nilai ralat yang diperolehi melalui MA(1) ialah 7.9 mikroterikan (28%), dan taburannya ditunjukkan dalam Rajah 3c. Berdasarkan kepada analisis kriteria penentuan purata-tergerak ini, ia menunjukkan bahawa pendekatan yang optimum ialah model MA(1), di mana ia menghasilkan nilai ralat yang terendah.

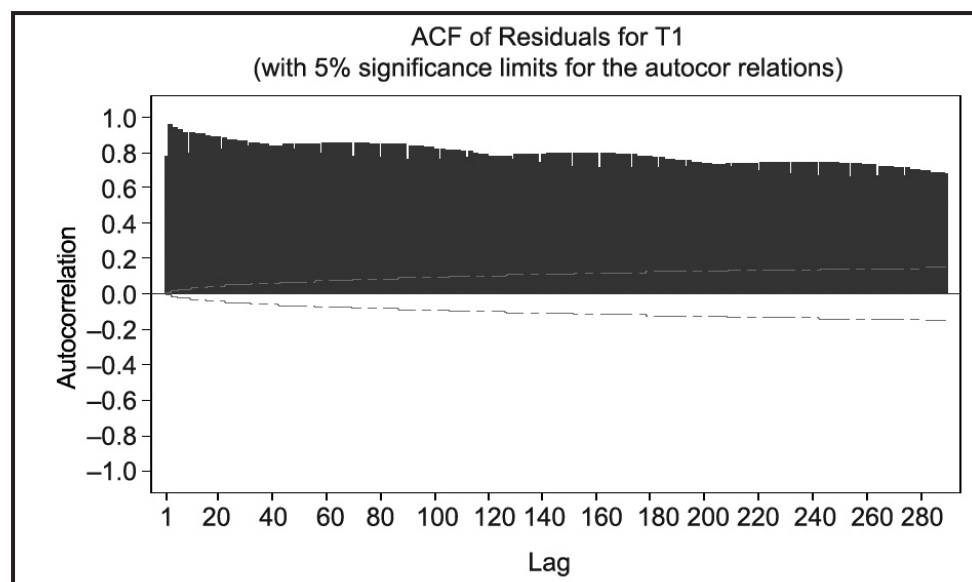


Rajah 3. Carta pai bagi purata berganda berdasarkan kepada kaedah: (a) MAPE, (b) MAD, (c) MSD

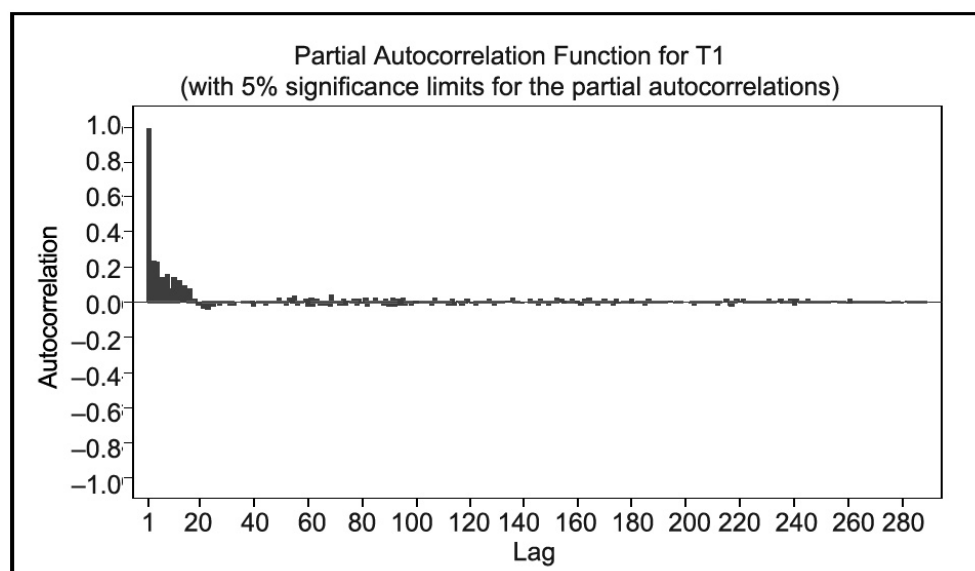
Rajah 4 menunjukkan taburan ACF (Rajah 4a) dan PACF (Rajah 4b) bagi model MA(1). Ia menunjukkan bahawa saling bergantung antara setiap sampel data ialah 93.5%. Taburan ini juga menunjukkan bahawa kewujudan satu corak pereputan yang berterusan dan ia menunjukkan terdapat banyak nilai-nilai separa autokorelasi yang bukan sifar. Keadaan ini dapat diperhatikan menerusi taburan plot dalam Rajah 4. Oleh itu, berdasarkan kepada Rajah 4, model ini menyarankan supaya set data terikan lesu ini

perlu dianalisis sebanyak 15 kali secara purata-tergerak bagi menghasilkan data yang stabil dan pegun.

Rajah 5a dan 5b masing-masing menunjukkan taburan statistik korelasi untuk pendekatan ACF dan PACF untuk model MA(2). Menerusi data analisis MA(2) ini, nilai θ yang diperolehi ialah -0.920. Rajah ini juga menunjukkan bahawa nilai parameter saling bergantung antara setiap yang dikaji ialah 92.0%.



(a)



(b)

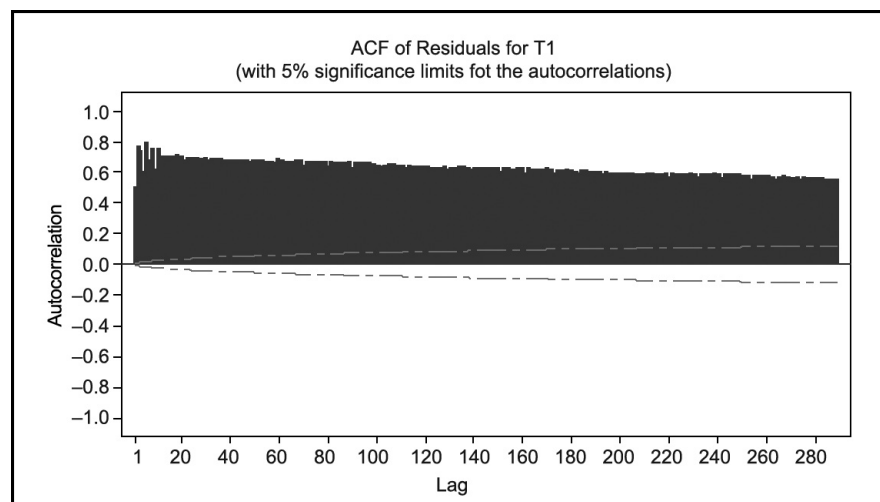
Rajah 4. (a) ACF untuk MA(1), (b) PACF untuk for MA(1)

Taburan tersebut menunjukkan terdapat satu corak pereputan yang berterusan dan juga terdapat banyak nilai-nilai separa autokorelasi yang bukan sifar dapat diperolehi dari taburan plot ini. Oleh itu, berdasarkan kepada Rajah 5, hasil daripada model ini menunjukkan bahawa set data tersebut perlu dianalisis sebanyak 15 kali secara purata-tergerak bagi menghasilkan data yang stabil dan pegun.

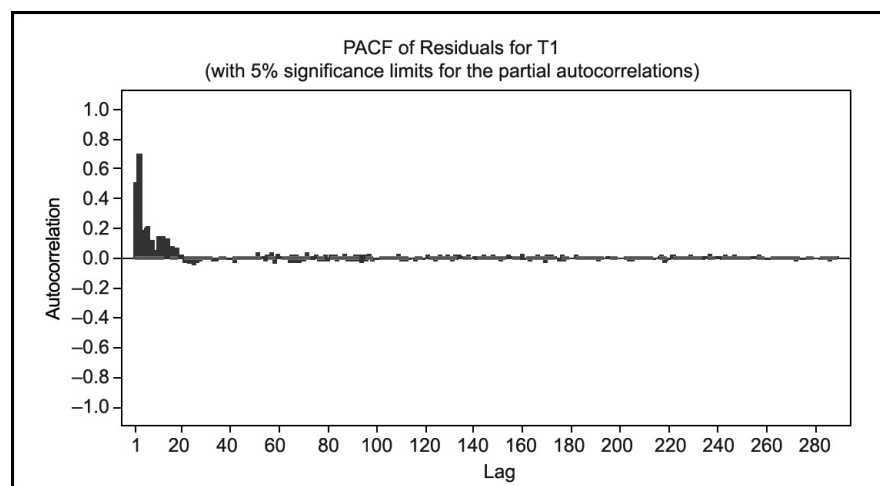
Rajah 6 menunjukkan taburan statistik korelasi untuk komponen ACF (Rajah 6a) dan PACF (Rajah 6b) bagi model MA(3). Sementara itu, analisis ini memberikan nilai θ iaitu -0.801. Keputusan yang dapat diperolehi daripada rajah ini menunjukkan bahawa saling bergantung

antara data tersebut ialah 80.1%. Berdasarkan kepada Rajah 6, Model MA(3) ini menyarankan supaya data set ini perlu dianalisis sebanyak 19 kali secara purata-tergerak bagi menghasilkan data yang stabil dan pegun.

Berdasarkan pendekatan-pendekatan pada analisis autokorelasi, dapat diperhatikan bahawa MA(1) menunjukkan korelasi yang paling tinggi berbanding MA(2) dan MA(3). Nilai autokorelasi bagi MA(1), MA(2) dan MA(3) masing-masing adalah 0.935, 0.920 dan 0.801. Oleh yang demikian, model MA(1) memberikan nilai saling kaitan yang lebih tinggi berbanding dengan MA(2) dan MA(3).

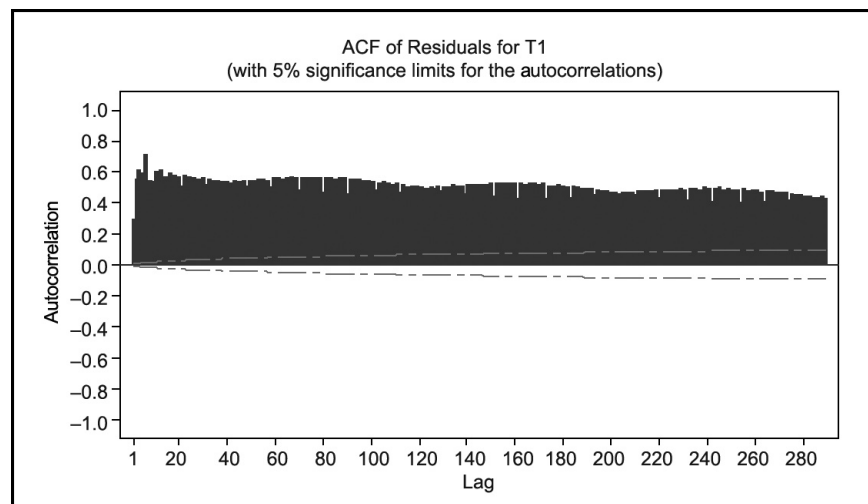


(a)

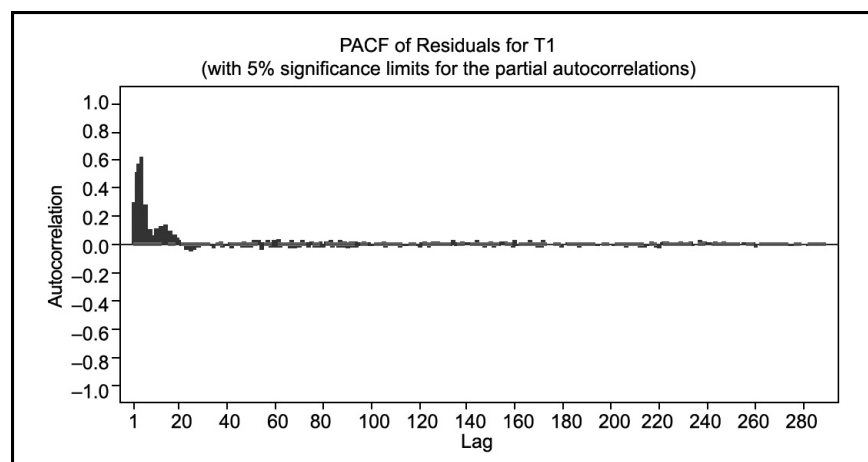


(b)

Rajah 5. (a) ACF bagi MA(2), (b) PACF bagi MA(2)



(a)



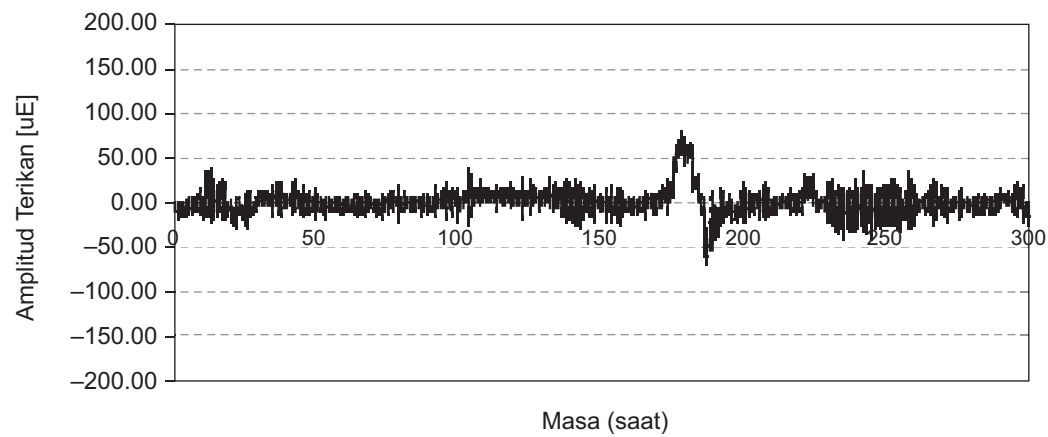
(b)

Rajah 6. (a) ACF bagi MA(3), (b) PACF bagi MA(3)

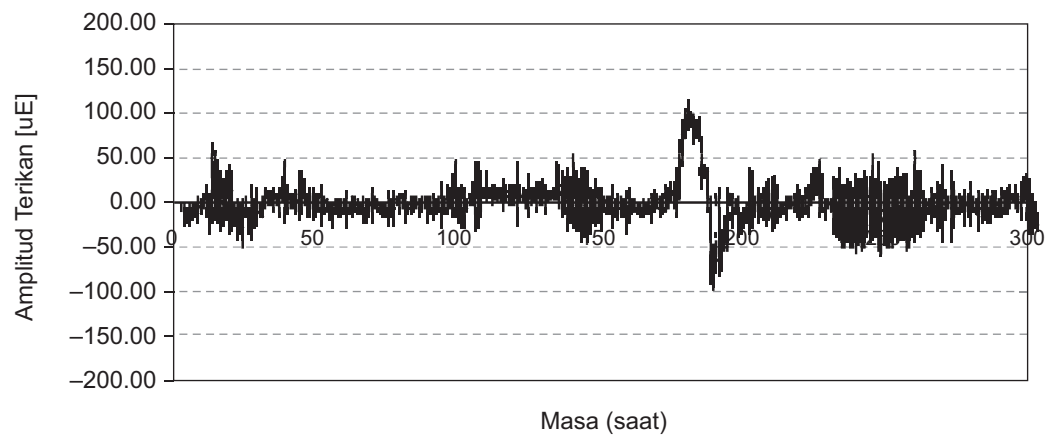
Menerusi Rajah 7 hingga Rajah 9, dapat ditunjukkan plot data siri-masa terikan lesu yang terhasil daripada ketiga-tiga model purata-tergerak, iaitu MA(1), MA(2) dan MA(3). Nilai-nilai statistik data yang terhasil ditunjukkan menerusi Jadual 2 yang mana perbandingan dengan nilai-nilai statistik asal boleh dilakukan. Jadual 3 menunjukkan perbezaan antara nilai-nilai statistik untuk data asal terikan-lesu dengan data yang dibentuk daripada aplikasi ketiga-tiga model purata-tergerak.

Dengan merujuk kepada Jadual 2, model MA(1) dapat mengurangkan kesan amplitud terikan dan menghasilkan isyarat baru yang agak pegun.

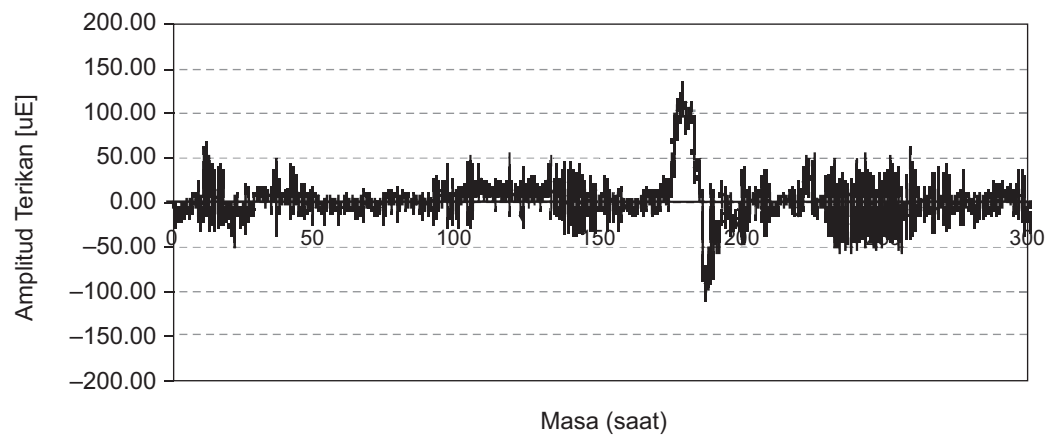
Ini dapat diperhatikan menerusi nilai kurtosis yang paling rendah jika dibandingkan dengan model MA(2) dan MA(3), iaitu pada aras 10.739. Di samping itu, perbezaan kurtosis antara model MA(1) dan data asal adalah tertinggi iaitu pada 21.9%. Ini menunjukkan kebanyakan data diskret yang menyebabkan kesan terlampau (outlier) dapat disingkirkan. Justeru, data siri-masa tidak pegun yang terbaik dapat dihasilkan menerusi model MA(1) berbanding dengan model MA(2) dan MA(3). Tambahan lagi, nilai minimum bagi data lesu ini ialah -138.72 mikroterikan dan nilai maksimum ialah 167.13 mikroterikan.



Rajah 7. Plot siri-masa untuk data asal terikan lesu pelbagai amplitud yang terhasil daripada analisis MA(1)



Rajah 8. Plot siri-masa untuk data asal terikan lesu pelbagai amplitud yang terhasil daripada analisis MA(2)



Rajah 9. Plot siri-masa untuk data asal terikan lesu pelbagai amplitud yang terhasil daripada analisis MA(3)

Jadual2. Nilai-nilai statistik data siri-masa terikan lesu yang terhasil daripada mode-model purata-tergerak

Keterangan	Data Asal	MA(1)	MA(2)	MA(3)
Min [$\mu\epsilon$]	2.337	2.337	2.337	2.337
Sisihan Piawai [$\mu\epsilon$]	25.466	12.356	17.500	19.946
Minimum [$\mu\epsilon$]	-138.720	-67.102	-92.229	-108.408
Maksimun [$\mu\epsilon$]	167.130	78.854	113.970	132.785
Kepencongan	1.869	1.520	1.649	1.678
Kurtosis	13.745	10.739	11.952	12.116

Jadual 3. Perbezaan nilai-nilai statistik data asal dengan data yang diperbaiki dengan kaedah purata-tergerak

Penerangan	Perbezaan [%]		
	Data Asal & MA(1)	Data Asal & MA(2)	Data Asal & MA(3)
Min (Purata)	0.0	0.0	0.0
Sisihan Piawai	51.5	31.3	21.7
Minimum	51.6	33.5	21.9
Maksimum	52.8	31.8	20.5
Kepencongan	18.7	11.8	10.2
Kurtosis	21.9	13.0	11.9

KESIMPULAN

Kajian ini dijalankan untuk menganalisis data terikan-lesu yang diukur pada sistem penyerap hentakan kereta. Keseluruhannya, kajian ini lebih fokus kepada pendekatan analisis data rawak terikan-lesu dengan menggunakan kaedah statistik, atau lebih spesifik kepada pendekatan kaedah purata-tergerak. Dalam kajian ini, data terikan-lesu telah diperolehi daripada pengukuran ke atas komponen sistem ampaian kereta yang dipandu di atas permukaan jalanraya luar kawasan perbandaran. Data ini diukur pada frekuensi persampelan 200 Hz dalam tempoh 300 saat.

Menerusi dilakukan analisis purata-tergerak menerusi kaedah Peratusan Ralat Min Mutlak (MAPE), Sisihan Min Mutlak (MAD) dan Sisihan Min Kuasa Dua (MSD), didapati model MA(1) lebih sesuai digunakan dalam analisis penghasilan corak data yang pegun berbanding dengan model MA(2) dan MA(3). Dalam analisis ini, MA(1) memberikan nilai 28%, 31% dan 33% masing-masing untuk kaedah MSD, MAD dan MAPE.

Analisis korelasi data berasaskan kepada Fungsi Autokorelasi (ACF) dan Fungsi Separa Autokorelasi (PACF) juga diperkenalkan dan dibincangkan. Menerusi aplikasi kedua-dua kaedah korelasi ini, nilai autokorelasi bagi MA(1)

(93.5%) adalah lebih tinggi berbanding dengan nilai MA(2) (92.0%) dan MA(3) (80.1%). Ini menunjukkan bahawa model MA(1) memberikan nilai saling kaitan yang lebih tinggi berbanding dengan MA(2) dan MA(3).

Akhirnya, dapat dirumuskan bahawa kaedah purata-tergerak ini dapat digunakan untuk memproses data rawak terikan-lesu yang berubah-ubah bagi penghasilan satu data yang pegun. Di samping itu, telah dikenalpasti yang model purata-tergerak MA(1) adalah yang terbaik jika dibandingkan dengan model MA(2) dan MA(3) dalam analisis rawak terikan-lesu ini. Ini dapat dibuktikan dengan nilai kurtosis terendah dapat dihasilkan pada nilai 10.739 (perbezaan tertinggi pada 21.9%), dan ini menunjukkan bahawa terdapat banyak data diskret yang terlampau (outlier) dapat disingkirkan bagi tujuan penghasilan data yang lebih baik dan bercorak lebih pegun.

PENGHARGAAN

Penulis-penulis merakamkan ucapan terima kasih kepada Kementerian Pengajian Tinggi (KPT) Malaysia dan Universiti Kebangsaan Malaysia atas sokongan yang diberikan dalam kajian ini menerusi peruntukan UKM-KK-02-FRGS0016.

RUJUKAN

- Agrawal, D. and Aggarwal, C. C. 2001. On the design and quantify of privacy preserving data mining algorithms. *Proc. of the 20th ACM SIGMOD Symposium on Principles of Database Systems*, Santa Barbara, 247-255.
- Agrawal, R. and Srikant, R. 2000. Privacy-preserving data mining. *Proc. of the ACM SIGMOD Conference on Management of Data*, Texas, USA, 439-450.
- Bendat, J. S. and Piersol, A. G. 1986. *Random Data: Analysis and Measurement Procedures*, New York: Wiley-Interscience.
- Estivill-Castro, V. and Brankovic, L. 1999. Data swapping: Balancing privacy against precision in mining for logic rules. *Proc. of the 1st Conference on Data Warehousing and Knowledge Discovery (DaWaK-99)*, Florence, Italy, 389-398.
- Harvey, A.C. 1981. *Time Series Model*. Oxford: Philip Allan Publishers Limited.

- Hinton, P.R. 1995. *Statistics Explained: A Guide for Social Science Students*, London: Routledge.
- Holt, C.C., Modigliani, F., Muth, J.F. and Simon, H.A. 1960. *Planning Production Inventories and Work Force*. New Jersey: Prentice Hall.
- Kendall, M.G., Stuart, A. and Ord, K. 1983. *The Advanced Theory of Statistics*. London: Charles Griffin.
- Makridakis, S., Wheelwright, S. C. and Hyndman, R. J. 1998. *Forecasting Methods and Applications*. 3rd Edition, New York: John Wiley & Sons, Inc.
- Meyer, Y. 1993. *Wavelets: Algorithm and Applications*. Philadelphia: SIAM.
- Natrella, M.G. 1966. *Experimental Statistics*, Handbook 91, Washington: NIST.
- Newland, D. E. 1993. *An Introduction to Random Vibrations Spectral and Wavelet Analysis*. 3rd Edition, New York: Longman Scientific and Technical.
- Qu, L. and He, Z. 1986. *Mechanical Diagnostics*, Shanghai: Shanghai Science and Technology Press.
- Tacer, B. and Loughlin, P.J. 1998. Nonstationary signal classification using the joint moments of time-frequency distributions, *Pattern Recognition*, 31: 1635-1641.
- Yan, J.H., Zheng, X. L. and Zhao, K. 2001. Experimental investigation on the small-load-omitting criterion. *Int. J. Fatigue*, 23: 403-415.